

Geospatial Data Exchange Protocol



Prepared By: Enterprise GIS Consortium
Date of Publication: 5/24/2011

Revision History

Version	Date	Author(s)	Comments
4a	11/26/2008	EGC	Approved by EGC
5a	5/31/2011	EGC	Approved by EGC via E-vote (email)

History: Vermont's Enterprise GIS Strategic Plan identifies several strategic objectives¹ oriented toward the sharing and exchange of geospatial data. Since 2008 members of the Enterprise GIS Consortium (EGC) have been sharing and exchanging data using the protocols established in the first version of this document. State agencies also continue to use a number of adhoc and informal methods to share and exchange geospatial data.

Several agencies have deployed ArcSDE (ESRI's² Spatial Database Engine) enabling the storage of geospatial data within a Relational Database Management System (RDBMS). Others have chosen to manage their data using a number of file-based storage solutions (eg: shapefiles, file geodatabase). Some agencies have established direct "links"³ to the Distributed VGIS Data Warehouse, allowing agencies and constituents throughout Vermont to easily find and download GIS data.

Back in 2006 VCGI and the State of Vermont drafted guidelines for the proper configuration and management of enterprise geodatabases (data stored in ArcSDE). One of the primary goals of the "ArcSDE Configuration and Management Guidelines"⁴ was to facilitate the exchange of data between Enterprise GIS servers (ArcSDE servers). Agencies responsible for managing ArcSDE servers (such as ACCD, ANR, VTrans, and VCGI) actively share and exchange data, and have already implemented the underlying architecture outlined in the ArcSDE guidelines. The exchange of data across an enterprise GIS has historically been a challenging endeavor. Exchanging and replicating data within an enterprise geodatabase (ArcSDE) environment adds additional requirements and challenges.

The Enterprise GIS Consortium (EGC) is responsible for the ongoing implementation and management of the Vermont's Enterprise GIS. The EGC is guided by the vision, goals, and strategies outlined in the Enterprise GIS Strategic Plan. The establishment and maintenance of a formal geodatabase exchange protocol has been identified as a high priority by the EGC.

Guiding Principals: The data exchange protocol must

1. Ensure consistency and predictability.
2. Be based on established standards.
3. Ensure data integrity through the use of appropriate data management practices and exchange mechanisms.
4. Allow for the exchange of information between heterogeneous environments.
5. Control access to "secure and confidential" data.

¹ Refer to Objective 10 and Objective 12.

² Environmental Systems Research Institute

³ VCGI has implemented a distributed data distribution architecture which provides users with a single portal for discovering and downloading Vermont's GIS (VGIS) data. This system is referred to as the Distributed VGIS Data Warehouse (Warehouse). The Warehouse consists of several components, including GIS data, a cataloging database, metadata, and data extract services. The Warehouse is administered by VCGI, but is supported by organizations who create, maintain, and provide GIS data (eg: ANR, VTrans, etc.).

⁴ http://www.vcgi.org/techres/standards/partiii_section_m.pdf

Scope: This data exchange protocol addresses the exchange of file-based and ArcSDE datasets. It addresses the sharing and exchange of data in the public domain⁵, as well as data that may have access constraints (confidential or sensitive data exempt from Federal and State public record law).

EGC Membership: Agencies and departments who want to share and receive data via this protocol must first become members of the Enterprise GIS Consortium (EGC).

Roles: Participants in the data exchange process fit into the following roles

1. **Publishers:** Agencies and departments that “publish” geospatial data. Publishers must comply with established data management standards and protocols, and must provide data in a manner consistent with the specifications outlined herein.
2. **Subscribers:** Agencies and departments that “subscribe” and receive geospatial data provided by Publishers. EGC membership is the only requirement.

Participants may choose to be Publishers, Subscribers, or both.

Data Management Standards: The EGC has drafted common database maintenance standards, procedures, and protocols that ensure consistency and predictability. Publishers are responsible for ensuring that their database’s (eg: GDB_VCGI in the case of VCGI) comply with these standards. Publishers of non-secure data must also be linked to the “Distributed VGIS Data Warehouse” system⁶, which help to ensure consistency over time. These requirements are limited to those databases that are shared and exchanged within Vermont’s Enterprise GIS environment.

Service Level Agreement: EGC members who choose to participate as "Publishers" are not required to provide any service level guarantee. However, "Publishers" are asked to notify "Subscribers" 48 hours prior to any planned system maintenance or down time.

Software Requirements: Most of the data exchange methods outline within this protocol require the use of ESRI GIS software.

Secure Data Exchange Requirements and Specifications: Secure and confidential is defined as one or more

- Personally identifiable information (PII) as defined by the State of Vermont.⁷
- Public record exemptions as defined by Federal and/or State Public Record Laws.⁸
- Other data which agencies or departments consider secure and confidential.

Requirements

⁵ Vermont public record law specifies that GIS data is a public record. State public record law defines specific exemptions; most GIS data is not exempt. However, state law requires that the state ensure that all GIS data “are stored and distributed in a manner which will limit the disclosure of data containing individual identifiers to disclosure consented to by the individuals in the data.” (Vermont Statute (Title 10 VSA - Chapter 8 § 121)).

⁶ http://www.vcgi.org/dataware/?page=../search_tools/linked_servers.cfm

⁷ <http://dii.vermont.gov/sites/dii/files/pdfs/Incident-Response-Policy.pdf>

⁸ <http://vermont-archives.org/records/handbook/pdf/PublicRecordsLaw.pdf>

- Secure/confidential data should be cataloged to help agencies identify specific datasets flagged as "secure and confidential".
- Secure/confidential data shall not be exchanged until a "Secure Data Exchange Memorandum of Understanding" has been signed by all parties involved in the exchange process.
- Secure/confidential data shall be exchanged as defined in the "Secure Data Exchange Memorandum of Understanding".
- Secure/confidential data shall be protected from unwarranted access as defined in the "Secure Data Exchange Memorandum of Understanding".

Specifications

- Secure/confidential data must
 - Comply with the "secure and confidential" naming convention for 1) Portable Data Packages, 2) Direct Data Connections, and 3) Geodata Services.
 - Comply with the Data Exchange Protocol's data management standards, which includes feature dataset/class naming conventions.

Methods: There are a number of mechanisms that can be used to exchange data between agencies. This protocol outlines three primary mechanisms; 1) Portable Data Packages, 2) Direct Data Connections, and 3) Geodata Services.

Portable Data Packages

Data Exchange via File Geodatabase: The most simplistic way of exchanging data is to use import/export methods. An ESRI ArcGIS File Geodatabase (FGDB) can be used as the "data exchange container". ArcGIS can be used to transfer data from an Enterprise Geodatabase (EGDB) or file-based storage (eg: shapefile) environment into a FGDB (eg: GDB_VCGI.gdb). The FGDB can be transferred to another agency by copying it over a network or using standard digital media (eg: DVD). The data can then be transferred into the target database (eg: ArcSDE) using ArcGIS. The data exchange and loading routines can be automated, allowing import/export to be done during off-peak hours.

Databases to be Exchanged: The following geodatabases will be exchanged⁹. Additional databases will be added as needed (eg: VDH may want to participant by sharing GDB_VDH).

Database	Steward/Source
GDB_VCGI	VCGI
GDB_ANR	ANR
GDB_VTrans	VTrans
GDB_ACCD	ACCD

Database Naming Convention: The following naming convention will be used as prescribed in the "ArcSDE Configuration and Management Guidelines"¹⁰.

⁹ as defined in the section 3.4.2 to the "ArcSDE Configuration and Management Guidelines"

¹⁰ http://www.vcgi.org/techres/standards/partiii_section_m.pdf

- **Naming Convention:** GDB_<AGENCY_ACRONYM>
 - Example: GDB_VCGI

The naming convention also supports the notion of application specific geodatabases which agencies may want to share with other EGC members.

- **Naming Convention (application specific):** GDB_<AGENCY_ACRONYM>_<APP>
 - Example: GDB_VCGI_BBVT (BroadbandVT.org data)

The following naming convention should be used for "secure and confidential data".

- **Naming Convention:** GDB_SECURE_<AGENCY_ACRONYM>
 - Example: GDB_SECURE_VCGI
- **Naming Convention (application specific):** GDB_SECURE_<AGENCY_ACRONYM>_<APP>
 - Example: GDB_SECURE_VCGI_VirtualVT (*to support VirtualVT application*)

Data Exchange Format:

- ESRI File Geodatabase (version 9.3)
- Data exchange manifest (in PDF format)
 - Organization and operator who generated the File Geodatabase.
 - Date the File Geodatabase was created.
 - List of all new/updated data contained in the database.
 - GDB "validation" reports.

Data Exchange Frequency: Data will be exchanged bi-annually.

- First Monday in January of the calendar year
- First Monday in July of the calendar year

Data Flow: The data flow will be unidirectional from "source" to "target". The "source" is defined as the organization responsible for managing the database. For example, ANR is the steward of GDB_ANR and therefore would be considered the "source" of GDB_ANR. The "target" would be any organization that elected to participate in the data exchange.

Direct Data Connections

Direct Data Connections: EGC members participating in the data exchange protocol will have the option of connecting directly to ArcSDE databases over GovNet.

Direct Connect Enabled Servers and Databases: EGC members will be able to connect to the servers and databases outlined below over GovNet. Additional databases will be added as needed (eg: VDH may want to participant by sharing GDB_VDH).

Database	Server
GDB_VCGI	data.vcgi.org
GDB_ANR	
GDB_VTrans	

GDB_ACCD	
----------	--

Database Naming Convention: The following naming convention will be used as prescribed in the "ArcSDE Configuration and Management Guidelines"¹¹.

- **Naming Convention:** GDB_<AGENCY_ACRONYM>
 - Example: GDB_VCGI

The naming convention also supports the notion of application specific geodatabases which agencies may want to share with other EGC members.

- **Naming Convention (application specific):** GDB_<AGENCY_ACRONYM>_<APP>
 - Example: GDB_VCGI_BBVT (BroadbandVT.org data)

The following naming convention should be used for "secure and confidential data".

- **Naming Convention:** GDB_SECURE_<AGENCY_ACRONYM>
 - Example: GDB_SECURE_VCGI
- **Naming Convention (application specific):** GDB_SECURE_<AGENCY_ACRONYM>_<APP>
 - Example: GDB_SECURE_VCGI_VirtualVT (*to support VirtualVT application*)

Data Exchange Method(s): Direct connect to server and database over GovNet. This connection type is limited to ESRI applications (eg: ArcGIS Desktop, ArcGIS Server, etc.) which support SDE "direct connect" over port 1433.

Data Exchange Frequency: EGC members will be able to connect to these servers and databases at any time.

Geodata Services

Geodata Services: Geospatial data will be accessible via ArcGIS Server geodata services¹² using ESRI ArcGIS clients (eg: ArcGIS Desktop). These services will be hosted by ArcGIS Server. EGC members participating in the data exchange protocol will have the option of connecting directly to these services over GovNet. Providers of these services may also choose to allow Subscribers to access these services over the Internet.

Providers of Geodata Services: EGC members will be able to connect to services hosted by the following servers over GovNet. Additional servers and services will be added as needed.

Geodata Service	Server
GDS_VCGI	web.vcgi.org
GDS_ANR	
GDS_VTrans	
GDS_ACCD	

¹¹ http://www.vcgi.org/techres/standards/partiii_section_m.pdf

¹² http://edndoc.esri.com/arcobjects/9.2/NET_Server_Doc/manager/publishing/geodata_service.htm

Database Naming Convention: The following naming convention will be used.

- **Naming Convention:** GDS_<AGENCY_ACRONYM>
 - Example: GDS_VCGI

The naming convention also supports the notion of application specific geodata services which agencies want to share with other EGC members.

- **Naming Convention (application specific):** GDS_<AGENCY_ACRONYM>_<APP>
 - Example: GDS_VCGI_VirtualVT

The following naming convention should be used for "secure and confidential data".

- **Naming Convention:** GDS_SECURE_<AGENCY_ACRONYM>
 - Example: GDS_SECURE_VCGI
- **Naming Convention (application specific):** GDS_SECURE_<AGENCY_ACRONYM>_<APP>
 - Example: GDS_SECURE_VCGI_VirtualVT (*to support VirtualVT application*)

Data Exchange Method(s): Client applications will connect using supported ESRI clients.

Data Exchange Frequency: EGC members will be able to connect to these servers and services at any time.